

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

## (i) APPLICANT:

(A) NAME: Max-Planck-Gesellschaft zur Foerderung der  
Wissenschaften e.V.  
(B) STREET: none  
(C) CITY: Berlin  
(E) COUNTRY: DE  
(F) POSTAL CODE (ZIP): none

(A) NAME: CNRS  
(B) STREET: rue de la cardonille  
(C) CITY: Montpellier Cedex 05  
(E) COUNTRY: FR  
(F) POSTAL CODE (ZIP): 34094

(ii) TITLE OF INVENTION: Nucleic acid molecules coding for mammalian  
tumor suppressor proteins and methods for their isolation

(iii) NUMBER OF SEQUENCES: 17

## (iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

## (2) INFORMATION FOR SEQ ID NO: 1:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2790 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

## (ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 542..2545

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GAATTCGGGA GAGCAAGCGG GCATCTCCTG GGCGCCGTCA TGGCTGCTTA GGCTGCGCTG	60
CCTGCGGATC GCGGATCCGG GATCGGAGAT CTGACGGCGA CGCCTGAGTC CGGCTAGGGT	120
AGGTCTGGGT TGGAGTCTGT GCCTGCTTCC TTGGCGTGTG GTTGTTCTCTG CTTGATTGCT	180
TCAGCGTGCC ATCGGCTTCG TATTTGCATA GGAGTCAGAG GAGTTAATCT TGTCTCCTCG	240
AAGATAGACT CTCATGGTTT ATGATCCATC TCTGTGAGAA GACTTTATTT GTCTGTCTCT	300

009180-081500

TCTCACAGGT	TTGAGTCTTC	AGACTTCTAC	AGAACTCCAT	AATATCTGCC	TCACAGCTGG		360
CTTTCCTGCT	CTCACAGAAG	ATACCCAGCT	ATTGTGCTCT	GGATCTCTCC	TGGCTGCTAG		420
GCTGTAGCGC	TGCCTTTCTG	GAGTCAGGCT	GTAGTGACTC	CCCACCTTCT	TTCTGTCTGG		480
GCTTAAATGG	CACAGCAGTT	CCTCAGCACA	TCTGAAGAAG	AAAGTGTGAG	AACCAAAGGC		540
C ATG GCT CCA TTC CGC TGT CAA AAA TGT GGC AAG TCC TTC GTC ACC							586
Met Ala Pro Phe Arg Cys Gln Lys Cys Gly Lys Ser Phe Val Thr							
1	5			10	15		
CTG GAG AAG TTC ACC ATT CAC AAT TAT TCC CAC TCC AGG GAG CGC CCA							634
Leu Glu Lys Phe Thr Ile His Asn Tyr Ser His Ser Arg Glu Arg Pro							
	20			25	30		
TTC AAG TGC TCG AAG GCT GAG TGT GGC AAA GCC TTC GTC TCC AAG TAT							682
Phe Lys Cys Ser Lys Ala Glu Cys Gly Lys Ala Phe Val Ser Lys Tyr							
	35			40	45		
AAG CTG ATG AGA CAC ATG GCC ACA CAC TCG CCA CAG AAG ATT CAC CAG							730
Lys Leu Met Arg His Met Ala Thr His Ser Pro Gln Lys Ile His Gln							
	50			55	60		
TGT ACT CAC TGT GAG AAG ACA TTC AAC CGG AAG GAC CAC CTG AAG AAC							778
Cys Thr His Cys Glu Lys Thr Phe Asn Arg Lys Asp His Leu Lys Asn							
	65			70	75		
CAC CTC CAG ACC CAC GAT CCC AAC AAG ATC TCC TAC GCG TGT GAC GAT							826
His Leu Gln Thr His Asp Pro Asn Lys Ile Ser Tyr Ala Cys Asp Asp							
	80			85	90		95
TGC GGC AAG AAG TAC CAC ACC ATG CTG GGC TAC AAG AGG CAC CTG GCC							874
Cys Gly Lys Lys Tyr His Thr Met Leu Gly Tyr Lys Arg His Leu Ala							
	100			105	110		
CTG CAC TCG GCG AGC AAT GGC GAT CTC ACC TGT GGG GTG TGC ACC CTG							922
Leu His Ser Ala Ser Asn Gly Asp Leu Thr Cys Gly Val Cys Thr Leu							
	115			120	125		
GAG CTG GGG AGC ACC GAG GTC CTG CTG GAC CAC CTC AAG TCT CAC GCG							970
Glu Leu Gly Ser Thr Glu Val Leu Leu Asp His Leu Lys Ser His Ala							
	130			135	140		
GAA GAA AAG GCC AAC CAG GCA CCC AGG GAG AAG AAA TAC CAG TGC GAC							1018
Glu Glu Lys Ala Asn Gln Ala Pro Arg Glu Lys Lys Tyr Gln Cys Asp							
	145			150	155		
CAC TGT GAT AGA TGC TTC TAC ACC CGG AAA GAT GTG CGT CGC CAC CTG							1066
His Cys Asp Arg Cys Phe Tyr Thr Arg Lys Asp Val Arg Arg His Leu							
	160			165	170		175
GTG GTC CAC ACA GGA TGC AAG GAC TTC CTG TGT CAG TTC TGT GCC CAG							1114
Val Val His Thr Gly Cys Lys Asp Phe Leu Cys Gln Phe Cys Ala Gln							
	180			185	190		
AGA TTT GGG CGC AAA GAC CAC CTC ACT CGT CAC ACC AAG AAG ACC CAC							1162
Arg Phe Gly Arg Lys Asp His Leu Thr Arg His Thr Lys Lys Thr His							
	195			200	205		

TCC Ser	CAG Gln	GAG Glu 210	CTG Leu	ATG Met	CAA Gln	GAG Glu	AAT Asn 215	ATG Met	CAG Gln	GCA Ala	GGA Gly	GAT Asp 220	TAC Tyr	CAG Gln	AGC Ser	1210
AAT Asn 225	TTC Phe	CAA Gln	CTC Leu	ATT Ile	GCG Ala 230	CCT Pro	TCA Ser	ACT Thr	TCG Ser	TTC Phe 235	CAG Gln	ATA Ile	AAG Lys	GTT Val	GAT Asp	1258
CCC Pro 240	ATG Met	CCT Pro	CCT Pro	TTC Phe 245	CAG Gln	CTA Leu	GGA Gly	GCG Ala	GCT Ala 250	CCC Pro	GAG Glu	AAC Asn	GGG Gly	CTT Leu	GAT Asp 255	1306
GGT Gly	GGC Gly	TTG Leu	CCA Pro	CCC Pro 260	GAG Glu	GTT Val	CAT His	GGT Gly 265	CTA Leu	GTG Val	CTT Leu	GCT Ala	GCC Ala 270	CCA Pro	GAA Glu	1354
GAA Glu	GCT Ala	CCC Pro	CAA Gln 275	CCC Pro	ATG Met	CCG Pro	CCC Pro 280	TTG Leu	GAG Glu	CCT Pro	TTG Leu	GAG Glu	CCT Pro 285	TTG Leu	GAG Glu	1402
CCT Pro	TTG Leu	GAG Glu 290	CCT Pro	TTG Leu	GAG Glu	CCG Pro	ATG Met 295	CAG Gln	TCT Ser	TTG Leu	GAG Glu	CCT Pro 300	TTG Leu	CAG Gln	CCT Pro	1450
TTG Leu 305	GAG Glu	CCG Pro	ATG Met	CAG Gln	CCT Pro	TTG Leu 310	GAG Glu	CCA Pro	ATG Met	CAG Gln	CCT Pro 315	TTG Leu	GAG Glu	CCG Pro	ATG Met	1498
CAG Gln 320	CCT Pro	TTA Leu	GAG Glu	CCT Pro	TTG Leu 325	GAG Glu	CCT Pro	CTG Leu	GAG Glu	CCG Pro 330	ATG Met	CAG Gln	CCT Pro	TTG Leu	GAG Glu 335	1546
CCG Pro	ATG Met	CAG Gln	CCT Pro	TTG Leu 340	GAG Glu	CCT Pro	ATG Met	CAG Gln	CCA Pro 345	ATG Met	CTG Leu	CCA Pro	ATG Met	CAG Gln	CCA Pro 350	1594
ATG Met	CAG Gln	CCA Pro	ATG Met 355	CAG Gln	CCA Pro	ATG Met	CAG Gln 360	CCA Pro	ATG Met	CTG Leu	CCA Pro	ATG Met	CAG Gln 365	CCA Pro	ATG Met	1642
CTG Leu	CCA Pro	ATG Met 370	CAG Gln	CCA Pro	ATG Met	CAG Gln	CCA Pro 375	ATG Met	CAG Gln	CCA Pro	ATG Met	CTG Leu 380	CCA Pro	ATG Met	CCA Pro	1690
GAG Glu 385	CCG Pro	TCT Ser	TTC Phe	ACT Thr	CTG Leu	CAC His 390	CCT Pro	GGC Gly	GTA Val	GTT Val	CCC Pro 395	ACC Thr	TCT Ser	CCT Pro	CCC Pro	1738
CCA Pro 400	ATT Ile	ATT Ile	CTT Leu	CAG Gln 405	GAG Glu	CAT His	AAG Lys	TAT Tyr	AAT Asn	CCT Pro 410	GTT Val	CCT Pro	ACC Thr	TCA Ser	TAT Tyr 415	1786
GCC Ala	CCA Pro	TTT Phe	GTA Val	GGC Gly 420	ATG Met	CCC Pro	GTC Val	AAA Lys	GCA Ala 425	GAT Asp	GGC Gly	AAG Lys	GCC Ala	TTT Phe 430	TGC Cys	1834
AAC Asn	GTG Val	GGT Gly	TTC Phe	TTT Phe	GAG Glu	GAA Glu	TTT Phe	CCT Pro	CTG Leu	CAA Gln	GAG Glu	CCT Pro	CAG Gln	GCG Ala	CCT Pro	1882

435	440	445	
CTC AAG TTC AAC CCA TGT TTT GAG ATG CCT ATG GAG GGG TTT GGG AAA Leu Lys Phe Asn Pro Cys Phe Glu Met Pro Met Glu Gly Phe Gly Lys 450 455 460			1930
GTC ACC CTG TCC AAA GAG CTG CTG GTA GAT GCT GTG AAT ATA GCC ATT Val Thr Leu Ser Lys Glu Leu Leu Val Asp Ala Val Asn Ile Ala Ile 465 470 475			1978
CCT GCC TCT CTG GAG ATT TCC TCC CTA TTG GGG TTT TGG CAG CTC CCC Pro Ala Ser Leu Glu Ile Ser Ser Leu Leu Gly Phe Trp Gln Leu Pro 480 485 490 495			2026
CCT CCT ACT CCC CAG AAT GGC TTT GTG AAT AGC ACC ATC CCT GTG GGG Pro Pro Thr Pro Gln Asn Gly Phe Val Asn Ser Thr Ile Pro Val Gly 500 505 510			2074
CCT GGG GAG CCA CTG CCC CAT AGG ATA ACC TGT CTG GCG CAG CAG CAG Pro Gly Glu Pro Leu Pro His Arg Ile Thr Cys Leu Ala Gln Gln Gln 515 520 525			2122
CCA CCG CCA CTG CCG CCG CCA CCA CCG CTG CCA CTG CCA CAG CCA CTG Pro Pro Pro Leu Pro Pro Pro Pro Pro Leu Pro Leu Pro Gln Pro Leu 530 535 540			2170
CCA GTG CCA CAG CCA CTA CCA CAG CCA CAG ATG CAG CCA CAG TTT CAG Pro Val Pro Gln Pro Leu Pro Gln Pro Gln Met Gln Pro Gln Phe Gln 545 550 555			2218
TTG CAG ATC CAG CCC CAG ATG CAG CTA CCA CAG CTG CTG CCG CAA CTG Leu Gln Ile Gln Pro Gln Met Gln Leu Pro Gln Leu Leu Pro Gln Leu 560 565 570 575			2266
CAA CCT CAG CAG CAG COT GAT CCT GAG CCA GAG CCA GAG CCA GAG CCA Gln Pro Gln Gln Gln Pro Asp Pro Glu Pro Glu Pro Glu Pro Glu Pro 580 585 590			2314
GAG CCA GAG CCA GAG CCA GAG CCG GAA CCG GAA CCG GAG CCA GAG CCA Glu Pro Glu Pro Glu Pro Glu Pro Glu Pro Glu Pro Glu Pro Glu Pro 595 600 605			2362
GAG CCA GAA CCA GAG CCA GAG GAA GAA CAG GAA GAG GCA GAA GAA GAG Glu Pro Glu Pro Glu Pro Glu Glu Glu Gln Glu Glu Ala Glu Glu Glu 610 615 620			2410
GCA GAG GAA GGA GCA GAG GAA GGA GCA GAA CCA GAG GCA CAG GCA GAA Ala Glu Glu Gly Ala Glu Glu Gly Ala Glu Pro Glu Ala Gln Ala Glu 625 630 635			2458
GAA GAG GAA GAG GAA GAG GAA GCG GAA GAG CCA CAG CCA GAA GAA GCC Glu Glu Glu Glu Glu Glu Glu Ala Glu Glu Pro Gln Pro Glu Glu Ala 640 645 650 655			2506
CAA ATA GCA GGA CTC GTC TAT AAG AAA TGG ACA GTT TAG TTCCTCTCT Gln Ile Ala Gly Leu Val Tyr Lys Lys Trp Thr Val *			2555
660 665			
TGTTAGCTTA CTCTGTAGTT TCTTCTTCTT GTTGCCCATTT GTGTAGCTTT ATAGAGTGTG			2615

003430-0345200

(2) INFORMATION FOR SEQ ID NO: 2:

(A) LENGTH: 668 amino acids

(D) TOPOLOGY: linear

```
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
```

Met 1	Ala	Pro	Phe	Arg 5	Cys	Gln	Lys	Cys	Gly 10	Lys	Ser	Phe	Val	Thr 15	Leu
Glu	Lys	Phe	Thr 20	Ile	His	Asn	Tyr	Ser 25	His	Ser	Arg	Glu	Arg 30	Pro	Phe
Lys	Cys	Ser 35	Lys	Ala	Glu	Cys	Gly 40	Lys	Ala	Phe	Val	Ser 45	Lys	Tyr	Lys
Leu	Met 50	Arg	His	Met	Ala	Thr 55	His	Ser	Pro	Gln	Lys 60	Ile	His	Gln	Cys
Thr 65	His	Cys	Glu	Lys	Thr 70	Phe	Asn	Arg	Lys	Asp 75	His	Leu	Lys	Asn	His 80
Leu	Gln	Thr	His	Asp 85	Pro	Asn	Lys	Ile	Ser 90	Tyr	Ala	Cys	Asp	Asp 95	Cys
Gly	Lys	Lys	Tyr 100	His	Thr	Met	Leu	Gly 105	Tyr	Lys	Arg	His	Leu 110	Ala	Leu
His	Ser	Ala 115	Ser	Asn	Gly	Asp	Leu 120	Thr	Cys	Gly	Val	Cys 125	Thr	Leu	Glu
Leu	Gly 130	Ser	Thr	Glu	Val	Leu 135	Leu	Asp	His	Leu	Lys 140	Ser	His	Ala	Glu
Glu 145	Lys	Ala	Asn	Gln	Ala 150	Pro	Arg	Glu	Lys	Lys 155	Tyr	Gln	Cys	Asp	His 160
Cys	Asp	Arg	Cys	Phe 165	Tyr	Thr	Arg	Lys	Asp 170	Val	Arg	Arg	His	Leu 175	Val
Val	His	Thr	Gly 180	Cys	Lys	Asp	Phe	Leu 185	Cys	Gln	Phe	Cys	Ala 190	Gln	Arg
Phe	Gly 195	Arg	Lys	Asp	His	Leu	Thr 200	Arg	His	Thr	Lys	Lys 205	Thr	His	Ser
Gln	Glu	Leu	Met	Gln	Glu	Asn	Met	Gln	Ala	Gly	Asp	Tyr	Gln	Ser	Asn

210	215	220
Phe Gln Leu Ile Ala Pro	Ser Thr Ser Phe Gln Ile Lys Val Asp Pro	
225	230	235 240
Met Pro Pro Phe Gln Leu Gly Ala Ala	Pro Glu Asn Gly Leu Asp Gly	
	245	250 255
Gly Leu Pro Pro Glu Val His Gly Leu Val Leu Ala Ala	Pro Glu Glu	
	260	265 270
Ala Pro Gln Pro Met Pro Pro	Leu Glu Pro Leu Glu Pro Leu Glu Pro	
	275	280 285
Leu Glu Pro Leu Glu Pro Met Gln Ser Leu Glu Pro Leu Gln Pro Leu		
	290	295 300
Glu Pro Met Gln Pro Leu Glu Pro Met Gln Pro Leu Glu Pro Met Gln		
	305	310 315 320
Pro Leu Glu Pro Leu Glu Pro Leu Glu Pro Met Gln Pro Leu Glu Pro		
	325	330 335
Met Gln Pro Leu Glu Pro Met Gln Pro Met Leu Pro Met Gln Pro Met		
	340	345 350
Gln Pro Met Gln Pro Met Gln Pro Met Leu Pro Met Gln Pro Met Leu		
	355	360 365
Pro Met Gln Pro Met Gln Pro Met Gln Pro Met Leu Pro Met Pro Glu		
	370	375 380
Pro Ser Phe Thr Leu His Pro Gly Val Val Pro Thr Ser Pro Pro Pro		
	385	390 395 400
Ile Ile Leu Gln Glu His Lys Tyr Asn Pro Val Pro Thr Ser Tyr Ala		
	405	410 415
Pro Phe Val Gly Met Pro Val Lys Ala Asp Gly Lys Ala Phe Cys Asn		
	420	425 430
Val Gly Phe Phe Glu Glu Phe Pro Leu Gln Glu Pro Gln Ala Pro Leu		
	435	440 445
Lys Phe Asn Pro Cys Phe Glu Met Pro Met Glu Gly Phe Gly Lys Val		
	450	455 460
Thr Leu Ser Lys Glu Leu Leu Val Asp Ala Val Asn Ile Ala Ile Pro		
	465	470 475 480
Ala Ser Leu Glu Ile Ser Ser Leu Leu Gly Phe Trp Gln Leu Pro Pro		
	485	490 495
Pro Thr Pro Gln Asn Gly Phe Val Asn Ser Thr Ile Pro Val Gly Pro		
	500	505 510
Gly Glu Pro Leu Pro His Arg Ile Thr Cys Leu Ala Gln Gln Gln Pro		
	515	520 525

00341370.024600

Pro Pro Leu Pro Pro Pro Pro Pro Leu Pro Leu Pro Gln Pro Leu Pro  
530 535 540

Val Pro Gln Pro Leu Pro Gln Pro Gln Met Gln Pro Gln Phe Gln Leu  
545 550 555 560

Gln Ile Gln Pro Gln Met Gln Leu Pro Gln Leu Leu Pro Gln Leu Gln  
565 570 575

Pro Gln Gln Gln Pro Asp Pro Glu Pro Glu Pro Glu Pro Glu Pro Glu  
580 585 590

Pro Glu Pro Glu Pro Glu Pro Glu Pro Glu Pro Glu Pro Glu Pro Glu  
595 600 605

Pro Glu Pro Glu Pro Glu Glu Glu Gln Glu Glu Ala Glu Glu Glu Ala  
610 615 620

Glu Glu Gly Ala Glu Glu Gly Ala Glu Pro Glu Ala Gln Ala Glu Glu  
625 630 635 640

Glu Glu Glu Glu Glu Glu Ala Glu Glu Pro Gln Pro Glu Glu Ala Gln  
645 650 655

Ile Ala Gly Leu Val Tyr Lys Lys Trp Thr Val \*

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

ATGTCTCGAG GCCTTTGCGG CCGCTATANN NNNNNN

36

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

0054370-031600

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

AGGCCATGG

9

(2) INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 9 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(iii) HYPOTHETICAL: YES

(ix) FEATURE:

- (A) NAME/KEY: -
- (B) LOCATION: 3
- (D) OTHER INFORMATION: /note= "N at position 3 is A or G."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

CCNCCATGG

9

(2) INFORMATION FOR SEQ ID NO: 6:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 9 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

His	Ser	Arg	Glu	Arg	Pro	Phe	Lys	Cys
1				5				

(2) INFORMATION FOR SEQ ID NO: 7:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 9 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(iii) HYPOTHETICAL: YES

0034870 02849200



## (ix) FEATURE:

- (A) NAME/KEY: Cross-links
- (B) LOCATION: 2
- (D) OTHER INFORMATION:/note= "X at position 2 is S or T."

## (ix) FEATURE:

- (A) NAME/KEY: Cross-links
- (B) LOCATION: 7
- (D) OTHER INFORMATION:/note= "X at position 7 is F or Y."

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

His Xaa Gly Glu Lys Pro Xaa Xaa Cys  
 1 5

## (2) INFORMATION FOR SEQ ID NO: 8:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (iii) HYPOTHETICAL: YES

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

His Ser Pro Gln Lys  
 1 5

## (2) INFORMATION FOR SEQ ID NO: 9:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (iii) HYPOTHETICAL: YES

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

Lys Lys Trp Thr  
 1

## (2) INFORMATION FOR SEQ ID NO: 10:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

003430 034300

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "oligonucleotid"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

GTGATGGCGG CCGCCATTCC GCTGTCAAAA ATGTG

35

(2) INFORMATION FOR SEQ ID NO: 11:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 30 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "oligonucleotide"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

CCGCGCCTCG AGGGTCTTCT TGGTGTGACG

30

(2) INFORMATION FOR SEQ ID NO: 12:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "oligonucleotide"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

GCGGCCGCAG AGCCGTCTTT CACTC

25

(2) INFORMATION FOR SEQ ID NO: 13:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 33 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "oligonucleotide"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

CCGCGCCTCG AGAACTGTCC ATTTCTTATA GAC

33

(2) INFORMATION FOR SEQ ID NO: 14:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs

009780 04095200

(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

ATAGCAGTGA GTGCTGTG

18

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 18 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

GTTTCTTTTC AGGGACTC

18

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 2334 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 803..2192

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

CGGTTCTTTC AATTCAGAAT TTGTTTTAGG TTCTGTTATT GCATAGATTT GCATACCTGT	60
TTTATGGTAT TTTAATACTG TTGGTTTTAA AAAATACCAT TTCCTCTGAG TGCTGTTCTG	120
AATATATTAT GTAAGCAATT TTGTGTGTTT TTTTTTTTCC ACTTGCATAA AGCAGGGGAA	180
AAGTTGAGAG TTTTTCTTAA TCCAGTCCCA AGTAGGACAA AGGATATGAG TGTTTAAAGA	240
TCATCTATTA AAATGCATGA AAAAACAATA GAAATCTCC TGTGCACATC GCCAGTCGTG	300
TGTGTGCTCT AGAAGTGAAG TTCAGGGGGT AACATAATGG AGGAATGTTT TCCTAGCTTC	360
ATTCCCTGAC GATGTACAAG GTCTCTTCTC ACAGGTTTGA ATCTTCAGAC AAACCTCTGG	420

009480-04500

GAGGACTGGG	AGGACTCGGT	CCCTGCCTCG	CAGCAGATGT	TCCCTGTAC	TCAGTAGCCA	480										
ATCCGGGGGA	CCCAGGACAT	GCCCCAGCTA	TAGTGATGCA	GATTACCTTT	CTGGTCCTGA	540										
ATCGCACCTG	TGCCTCGAGA	CTTTCTCCCC	TCAGCTTGAG	ACTGCATGTA	AACTGGGATG	600										
TGTGAAAGCA	GGAAGCAAAG	CTAGTGACAG	CTGAGAGGTC	CATGTCTGGG	TAGAACCAGG	660										
CCCACGATGC	TGCCTCTCCC	GTGGTCTGGA	GTTCAGCTGC	AGGGACTCTG	CTGATTGGCC	720										
CAGCACCATC	GTTCTGTTTG	TGCTTAAATG	GCACAGCATT	TGGTCAGCAC	ATCTGAAAAG	780										
GAAGGTGTGA	GAAGCAAAGC	CC	ATG	GCC	ACG	TTC	CCC	TGC	CAG	TTA	TGT	GGC	832			
			Met	Ala	Thr	Phe	Pro	Cys	Gln	Leu	Cys	Gly				
			1				5					10				
AAG	ACG	TTC	CTC	ACC	CTG	GAG	AAG	TTC	ACG	ATT	CAC	AAT	TAT	TCC	CAC	880
Lys	Thr	Phe	Leu	Thr	Leu	Glu	Lys	Phe	Thr	Ile	His	Asn	Tyr	Ser	His	
			15						20					25		
TCC	AGG	GAG	CGG	CCG	TAC	AAG	TGT	GTC	CAG	CCT	GAC	TGT	GGC	AAA	GCC	928
Ser	Arg	Glu	Arg	Pro	Tyr	Lys	Cys	Val	Gln	Pro	Asp	Cys	Gly	Lys	Ala	
			30					35					40			
TTT	GTT	TCC	AGA	TAT	AAA	TTG	ATG	AGG	CAT	ATG	GCT	ACC	CAT	TCT	CCC	976
Phe	Val	Ser	Arg	Tyr	Lys	Leu	Met	Arg	His	Met	Ala	Thr	His	Ser	Pro	
			45				50					55				
CAG	AAA	TCT	CAC	CAG	TGT	GCT	CAC	TGT	GAG	AAG	ACG	TTC	AAC	CGG	AAA	1024
Gln	Lys	Ser	His	Gln	Cys	Ala	His	Cys	Glu	Lys	Thr	Phe	Asn	Arg	Lys	
			60				65				70					
GAC	CAC	CTG	AAA	AAC	CAC	CTC	CAG	ACC	CAC	GAC	CCC	AAC	AAA	ATG	GCC	1072
Asp	His	Leu	Lys	Asn	His	Leu	Gln	Thr	His	Asp	Pro	Asn	Lys	Met	Ala	
			75			80				85					90	
TTT	GGG	TGT	GAG	GAG	TGT	GGG	AAG	AAG	TAC	AAC	ACC	ATG	CTG	GGC	TAT	1120
Phe	Gly	Cys	Glu	Glu	Cys	Gly	Lys	Lys	Tyr	Asn	Thr	Met	Leu	Gly	Tyr	
				95					100					105		
AAG	AGG	CAC	CTG	GCC	CTC	CAT	GCG	GCC	AGC	AGT	GGG	GAC	CTC	ACC	TGT	1168
Lys	Arg	His	Leu	Ala	Leu	His	Ala	Ala	Ser	Ser	Gly	Asp	Leu	Thr	Cys	
			110					115					120			
GGG	GTC	TGT	GCC	CTG	GAG	CTA	GGG	AGC	ACC	GAG	GTG	CTA	CTG	GAC	CAC	1216
Gly	Val	Cys	Ala	Leu	Glu	Leu	Gly	Ser	Thr	Glu	Val	Leu	Leu	Asp	His	
			125				130					135				
CTC	AAA	GCC	CAT	GCG	GAA	GAG	AAG	CCC	CCT	AGC	GGA	ACC	AAG	GAA	AAG	1264
Leu	Lys	Ala	His	Ala	Glu	Glu	Lys	Pro	Pro	Ser	Gly	Thr	Lys	Glu	Lys	
			140				145					150				
AAG	CAC	CAG	TGC	GAC	CAC	TGT	GAA	AGA	TGC	TTC	TAC	ACC	CGG	AAG	GAT	1312
Lys	His	Gln	Cys	Asp	His	Cys	Glu	Arg	Cys	Phe	Tyr	Thr	Arg	Lys	Asp	
					160					165					170	

GTG Val	CGA Arg	CGC Arg	CAC His	CTG Leu 175	GTG Val	GTC Val	CAC His	ACA Thr	GGA Gly 180	TGC Cys	AAG Lys	GAC Asp	TTC Phe	CTG Leu 185	TGC Cys	1360
CAG Gln	TTC Phe	TGT Cys	GCC Ala 190	CAG Gln	AGA Arg	TTT Phe	GGG Gly 195	CGC Arg	AAG Lys	GAT Asp	CAC His	CTC Leu	ACC Thr 200	CGG Arg	CAT His	1408
ACC Thr	AAG Lys	AAG Lys 205	ACC Thr	CAC His	TCA Ser	CAG Gln	GAG Glu 210	CTG Leu	ATG Met	AAA Lys	GAG Glu	AGC Ser 215	TTG Leu	CAG Gln	ACC Thr	1456
GGA Gly 220	GAC Asp	CTT Leu	CTG Leu	AGC Ser	ACC Thr	TTC Phe 225	CAC His	ACC Thr	ATC Ile	TCG Ser	CCT Pro 230	TCA Ser	TTC Phe	CAA Gln	CTG Leu	1504
AAG Lys 235	GCT Ala	GCT Ala	GCC Ala	TTG Leu	CCT Pro 240	CCT Pro	TTC Phe	CCT Pro	TTA Leu	GGA Gly 245	GCT Ala	TCT Ser	GCC Ala	CAG Gln	AAC Asn 250	1552
GGG Gly	CTT Leu	GCA Ala	AGT Ser	AGC Ser 255	TTG Leu	CCA Pro	GCT Ala	GAG Glu	GTC Val 260	CAT His	AGC Ser	CTC Leu	ACC Thr 265	CTC Leu	AGT Ser	1600
CCC Pro	CCA Pro	GAA Glu	CAA Gln 270	GCC Ala	GCC Ala	CAG Gln	CCT Pro 275	ATG Met	CAG Gln	CCG Pro	CTG Leu	CCA Pro	GAG Glu 280	TCC Ser	CTG Leu	1648
GCC Ala	TCC Ser	CTC Leu 285	CAC His	CCC Pro	TCG Ser	GTA Val	TCC Ser 290	CCT Pro	GGC Gly	TCT Ser	CCT Pro	CCG Pro 295	CCA Pro	CCC Pro	CTT Leu	1696
CCC Pro	AAT Asn 300	CAC His	AAG Lys	TAC Tyr	AAC Asn	ACC Thr 305	ACT Thr	TCT Ser	ACC Thr	TCA Ser	TAC Tyr 310	TCC Ser	CCA Pro	CTT Leu	GCA Ala	1744
AGC Ser 315	CTG Leu	CCC Pro	CTC Leu	AAA Lys 320	GCA Ala	GAT Asp	ACT Thr	AAA Lys	GGT Gly	TTT Phe 325	TGC Cys	AAT Asn	ATC Ile	AGT Ser	TTG Leu 330	1792
TTT Phe	GAG Glu	GAC Asp	TTG Leu 335	CCT Pro	CTG Leu	CAA Gln	GAG Glu	CCT Pro	CAG Gln 340	TCA Ser	CCT Pro	CAA Gln	AAG Lys 345	CTC Leu	AAC Asn	1840
CCA Pro	GGT Gly	TTT Phe	GAT Asp 350	CTG Leu	GCT Ala	AAG Lys	GGA Gly 355	AAT Asn	GCT Ala	GGT Gly	AAA Lys	GTA Val	AAC Asn 360	CTG Leu	CCC Pro	1888
AAG Lys	GAG Glu 365	CTG Leu	CCT Pro	GCA Ala	GAT Asp	GCT Ala	GTG Val 370	AAC Asn	CTA Leu	ACA Thr	ATA Ile 375	CCT Pro	GCC Ala	TCT Ser	CTG Leu	1936
GAC Asp	CTG Leu	TCC Ser	CCC Pro	CTG Leu	TTG Leu	GGC Gly	TTC Phe	TGG Trp	CAG Gln	CTG Leu	CCC Pro	CCT Pro	CCT Pro	GCT Ala	ACC Thr	1984

380	385	390	
CAA AAT ACC TTT GGG AAT AGC ACT CTT GCC CTG GGG CCT GGG GAA TCT			2032
Gln Asn Thr Phe Gly Asn Ser Thr Leu Ala Leu Gly Pro Gly Glu Ser			
395	400	405	410
TTG CCC CAC AGG TTA AGC TGT CTG GGG CAG CAG CAG CAA GAA CCC CCA			2080
Leu Pro His Arg Leu Ser Cys Leu Gly Gln Gln Gln Gln Glu Pro Pro			
	415	420	425
CTT GCC ATG GGC ACT GTG AGC CTG GGC CAG CTC CCC CTG CCC CCC ATC			2128
Leu Ala Met Gly Thr Val Ser Leu Gly Gln Leu Pro Leu Pro Pro Ile			
	430	435	440
CCT CAT GTG TTC TCA GCT GGC ACT GGC TCT GCC ATC CTG CCT CAT TTC			2176
Pro His Val Phe Ser Ala Gly Thr Gly Ser Ala Ile Leu Pro His Phe			
	445	450	455
CAT CAT GCA TTC AGA T AATTGATTTT TAAAGTGTAT TTTTCGTATT			2222
His His Ala Phe Arg			
	460		
CTGGAAGATG TTTTAAGAAG CATTTTAAAT GTCAGTTACA ATATGAGAAA GATTTGGAAA			2282
ACGAGACTGG GACTATGGCT TATTCAGTGA TGA CTGGCTT GAGATGATAA GA			2334

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 463 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

Met	Ala	Thr	Phe	Pro	Cys	Gln	Leu	Cys	Gly	Lys	Thr	Phe	Leu	Thr	Leu
1				5					10				15		
Glu	Lys	Phe	Thr	Ile	His	Asn	Tyr	Ser	His	Ser	Arg	Glu	Arg	Pro	Tyr
		20						25					30		
Lys	Cys	Val	Gln	Pro	Asp	Cys	Gly	Lys	Ala	Phe	Val	Ser	Arg	Tyr	Lys
		35					40					45			
Leu	Met	Arg	His	Met	Ala	Thr	His	Ser	Pro	Gln	Lys	Ser	His	Gln	Cys
	50					55					60				
Ala	His	Cys	Glu	Lys	Thr	Phe	Asn	Arg	Lys	Asp	His	Leu	Lys	Asn	His
	65				70					75					80
Leu	Gln	Thr	His	Asp	Pro	Asn	Lys	Met	Ala	Phe	Gly	Cys	Glu	Glu	Cys
				85					90					95	

00254970-081600

Gly Lys Lys Tyr Asn Thr Met Leu Gly Tyr Lys Arg His Leu Ala Leu  
 100 105 110  
 His Ala Ala Ser Ser Gly Asp Leu Thr Cys Gly Val Cys Ala Leu Glu  
 115 120 125  
 Leu Gly Ser Thr Glu Val Leu Leu Asp His Leu Lys Ala His Ala Glu  
 130 135 140  
 Glu Lys Pro Pro Ser Gly Thr Lys Glu Lys Lys His Gln Cys Asp His  
 145 150 155 160  
 Cys Glu Arg Cys Phe Tyr Thr Arg Lys Asp Val Arg Arg His Leu Val  
 165 170 175  
 Val His Thr Gly Cys Lys Asp Phe Leu Cys Gln Phe Cys Ala Gln Arg  
 180 185 190  
 Phe Gly Arg Lys Asp His Leu Thr Arg His Thr Lys Lys Thr His Ser  
 195 200 205  
 Gln Glu Leu Met Lys Glu Ser Leu Gln Thr Gly Asp Leu Leu Ser Thr  
 210 215 220  
 Phe His Thr Ile Ser Pro Ser Phe Gln Leu Lys Ala Ala Ala Leu Pro  
 225 230 235 240  
 Pro Phe Pro Leu Gly Ala Ser Ala Gln Asn Gly Leu Ala Ser Ser Leu  
 245 250 255  
 Pro Ala Glu Val His Ser Leu Thr Leu Ser Pro Pro Glu Gln Ala Ala  
 260 265 270  
 Gln Pro Met Gln Pro Leu Pro Glu Ser Leu Ala Ser Leu His Pro Ser  
 275 280 285  
 Val Ser Pro Gly Ser Pro Pro Pro Pro Leu Pro Asn His Lys Tyr Asn  
 290 295 300  
 Thr Thr Ser Thr Ser Tyr Ser Pro Leu Ala Ser Leu Pro Leu Lys Ala  
 305 310 315 320  
 Asp Thr Lys Gly Phe Cys Asn Ile Ser Leu Phe Glu Asp Leu Pro Leu  
 325 330 335  
 Gln Glu Pro Gln Ser Pro Gln Lys Leu Asn Pro Gly Phe Asp Leu Ala  
 340 345 350  
 Lys Gly Asn Ala Gly Lys Val Asn Leu Pro Lys Glu Leu Pro Ala Asp  
 355 360 365  
 Ala Val Asn Leu Thr Ile Pro Ala Ser Leu Asp Leu Ser Pro Leu Leu  
 370 375 380

00264370 034600

Gly Phe Trp Gln Leu Pro Pro Pro Ala Thr Gln Asn Thr Phe Gly Asn  
 385 390 395 400

Ser Thr Leu Ala Leu Gly Pro Gly Glu Ser Leu Pro His Arg Leu Ser  
 405 410 415

Cys Leu Gly Gln Gln Gln Gln Glu Pro Pro Leu Ala Met Gly Thr Val  
 420 425 430

Ser Leu Gly Gln Leu Pro Leu Pro Pro Ile Pro His Val Phe Ser Ala  
 435 440 445

Gly Thr Gly Ser Ala Ile Leu Pro His Phe His His Ala Phe Arg  
 450 455 460

005437 02845200